

科学研究費助成事業 研究成果報告書

平成 30 年 6 月 19 日現在

機関番号：82626

研究種目：基盤研究(C) (一般)

研究期間：2015～2017

課題番号：15K00328

研究課題名(和文) 機械学習と網羅シミュレーションによるMA共有資源選択の効率化・安定化手法の確立

研究課題名(英文) Improvement and Stabilization of Multiagent Resource Share Problem by Machine Learning and Exhaustive Simulation

研究代表者

野田 五十樹 (Noda, Itsuki)

国立研究開発法人産業技術総合研究所・情報・人間工学領域・総括研究主幹

研究者番号：40357744

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：ボトムアップアプローチとして、マルチエージェント資源共有問題のうち、特に資源の容量が動的に変化する動的資源共有問題を取り上げ、独立して強化学習を行うエージェントの探索率を学習を通じて調整する方法として、Win or Update Exploration-ratio (WoUE)を提案した。
トップダウンアプローチとしては、マルチエージェント資源共有問題のうち、特に大規模な人流を対象として取り上げて、トップダウンに人流を制御する場合の手法として、静的および動的な分断誘導を取り上げ、シミュレーションにより評価を行った。

研究成果の概要(英文)：As the bottom-up approach, we focused on multi-agent resource sharing problem with dynamically changing resources. We proposed a method called "Win or Update Exploration-ratio (WoUE)", in which each agent adjusts its exploration ratio through learning independently.

As the top-down approach, we picked up large scale pedestrian control at big events. We evaluated two top-down control methods, that are, fixed and adaptive dividing guidance of peoples via a simulator, and compared with real data.

研究分野：人工知能

キーワード：multiagent resource shareing learning simulation social simulation

1. 研究開始当初の背景

自律分散型のスマート社会実現のアプローチは、絶えず変化する社会システムや参加エージェントの多様さに対応しやすいという長所がある反面、個々のエージェントが利己的に最適なサービスを求めるあまり過度の集中が生じるというジレンマを内包している。例えば、渋滞や混雑の集中は、個々の人々における利得最大化が、交通という社会機能を損ない、ひいては各個人の効率を落としている。また、経済では、人々が同時に同じ品を売り/買いすることで、品不足や供給過剰、価格の高騰/下落が生じ、社会全体としての不具合が生じる。渋滞情報や価格変動情報といった情報提供によりこれらの不具合をある程度解消させることはできるが、過度の情報提供がさらなる集中・混乱を生じることも知られている。より進化したスマート社会の実現のためには、このようなジレンマ状態を解消する方法の確立が望まれる。

2. 研究の目的

マルチエージェントの共有資源選択問題でのジレンマについて、理論的及び実践的な手法で分析を行い、ジレンマ解消に向けた基盤技術の確立を目指す。スマート社会の実現の中央制御型と自律分散型の2つアプローチの内、自律分散型を取り上げ、エージェントの利己性に起因する過度集中のジレンマを共有資源選択問題として定式化し、エージェント行動選択の側面と、情報提供など群全体への誘導方策の側面の、両面からジレンマ解消に資する手法の研究開発を行う。提案内容は、学習エージェント群の挙動分析と網羅的シミュレーションによる誘導手法の評価など、提案者らのこれまでの取り組みに立脚しており、社会システム設計への新しい技術の切り口の提供が期待できる。

3. 研究の方法

マルチエージェント共有資源選択問題に対し、エージェントの行動様式を調整することで共有資源選択の効率化や安定化を目指すボトムアップアプローチと、マルチエージェント社会シミュレーションにより誘導方策などを用いた共有資源選択の集中回避等を目指すトップダウンアプローチの2つの方法で研究を進める。ボトムアップアプローチでは、これまで進めてきた動的環境におけるマルチエージェント強化学習の枠組みを拡張し、多様性を持つエージェント群での学習と群の挙動の関係を洗い出す。また、トップダウンアプローチでは、網羅的マルチエージェント社会シミュレーションを用いて、情報配信などの誘導方策とそれによる集中回避の効果の関係を、人流誘導などを題材に分析・可視化を行う。

4. 研究成果

4. 1. ボトムアップアプローチ

本研究では、図1に示すようなマルチエージェント資源共有問題のうち、特に資源の容量が動的に変化する動的資源共有問題を取り上げ、独立して強化学習を行うエージェントの探索率を学習を通じて調整する方法を提案した。

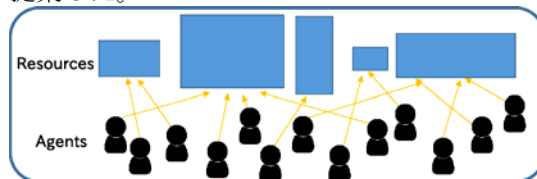


図1 マルチエージェント
資源共有問題

マルチエージェント学習においては、探索率は注意深く設定されなければならない。学習を促進するためには探索率を高く設定する必要があるが、多数のエージェントが探索を行うと、それはほかのエージェントの学習にとってノイズとなり、学習制度を悪化させる原因となる。一方で、動的環境下では環境の変化に追従するためにも、探索率を一定の値に保つ必要がある。従来の探索率調整の研究では、シングルエージェントあるいは静的環境を仮定したものがほとんどであり、上記のような問題設定では有効に働くものはなかった。

本研究ではこの問題に対して、Win or Update Exploration Ratio (WoUE) という手法を提案した。WoUEは、マルチエージェント学習の学習率を変化させる Win or Learn First (WoLF) の方法の考え方をベースにした探索率調整法となっている。考え方としては、自らの行動の成績がほかのエージェントの平均よりも悪い時にのみ、探索率を、成績を向上させる方向に変化させる。具体的なアルゴリズムを、図2に示す。

Algorithm 1 WoUE.

```

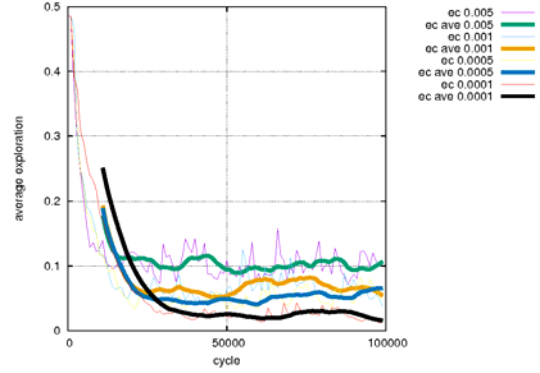
1: Initialize N agents
2: for cycle = 1 → endCycle do
3:   for all Agents do
4:     select action, using  $\epsilon$ -greedy
5:   end for
6:   evaluate all agent reward
7:   if interval MOD cycle = 0 then
8:     broadcast  $\mu_G$  to each agent
9:     for all Agents do
10:      if  $r_i(\epsilon_i(t)) < \mu_G$  then
11:        calculate  $\Delta\epsilon_i$  by Eq. (1)
12:         $\epsilon_i(t+1) \leftarrow \text{CHANGEBYSIGMOID}(\Delta\epsilon_i, \epsilon_i(t))$ 
13:         $\epsilon_i(t+1) \leftarrow \text{CHECKMINMAX}(\epsilon_i(t+1))$ 
14:      end if
15:    end for
16:  end if
17: end for
18:
19: function CHANGEBYSIGMOID( $\Delta\epsilon, \epsilon$ )
20:   sig_ $\epsilon \leftarrow \text{inverseSigmoid}(\epsilon)$ 
21:    $\epsilon \leftarrow \text{sigmoid}(\text{sig}_\epsilon + \Delta\epsilon)$ 
22:   return  $\epsilon$ 
23: end function
24:
25: function CHECKMINMAX( $\epsilon$ )
26:   if  $\epsilon < \text{minimum}_\epsilon$  then
27:      $\epsilon \leftarrow \text{minimum}_\epsilon$ 
28:   else if  $\text{maximum}_\epsilon < \epsilon$  then
29:      $\epsilon \leftarrow \text{maximum}_\epsilon$ 
30:   end if
31:   return  $\epsilon$ 
32: end function

```

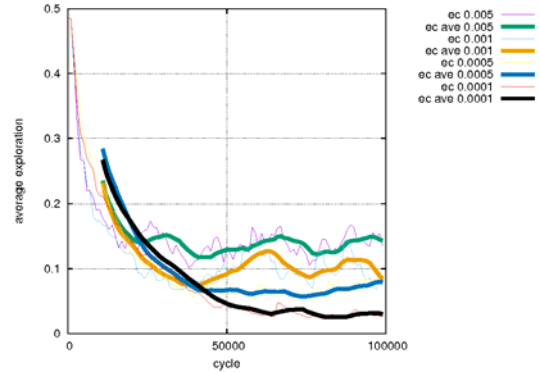
図 2 提案アルゴリズム

この WoUE は、エージェント集団全体としては、成績の分散を小さくさせる効果が数学的に確認されており、その結果、全体的な成績が向上することが見込まれる。

この提案手法を具体的な問題に適用し、効果を確認した実験結果を、図 3 に示す。実験の設定では、様々な変化率で容量が変化する資源が複数ある環境を用意し、そこにおいて、複数エージェント (50 エージェントおよび 100 エージェント) がうまく資源共有するよう学習を進める。グラフでは、その学習の過程で、各エージェントが獲得した探索率 ϵ の各サイクルでの平均と、さらにその移動平均を、環境の変化率別に示している。一般に、変化率の大きな環境ほど探索率も大きくとる必要がある。このグラフからは、変化率の大きさに応じて順に大きな探索率をエージェントが獲得していることが確認できる。



(a) 50 agents



(b) 100 agents

図 3 学習による探索率 ϵ の変化

さらに、獲得された探索率が最適の探索率にどれくらい対応できているかを比較したものが図 4 である。この図のうち、横軸の最適探索率は、WoUE を用いずに固定した探索率で学習を行うエージェントを用意し、各環境変化率に対してどの探索率が最適であるかをもとめたものである。一方、縦軸は、WoUE および従来手法 (MAER) により求められた探索率の平均である。この図から、WoUE で求められる探索率は最適探索率と強い相関を持つことがわかる。一方で、従来手法では最適探索率に対して低めの探索率になりがちであり、かつ、相関もばらけていることが読み取れる。このように、提案手法の WoUE は環境に応じた探索率を求めることができることが確認できる。

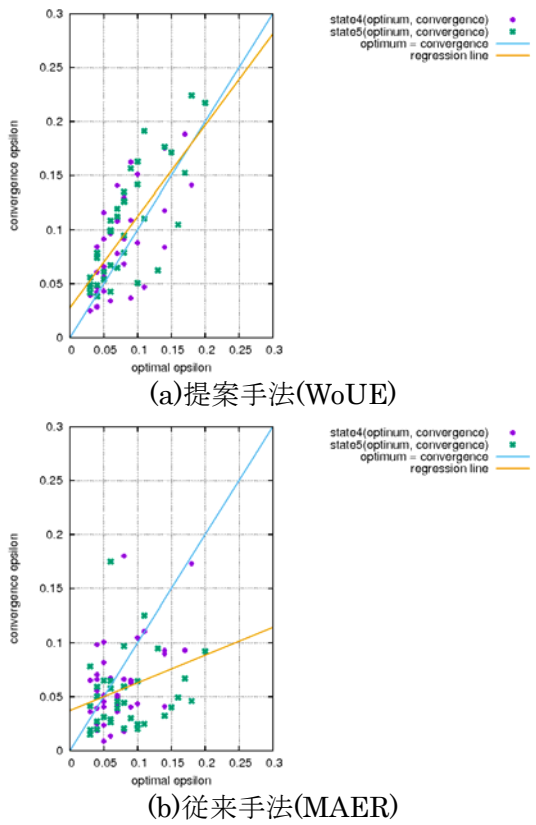


図 4 最適探索率との相関

ここで提案した WoUE は、状態遷移を含む強化学習課題である Dynamic Beach Problem Domain (DBPD, 図 5) にも適用され、同じく環境に適した探索率の獲得に成功している(図 6)。

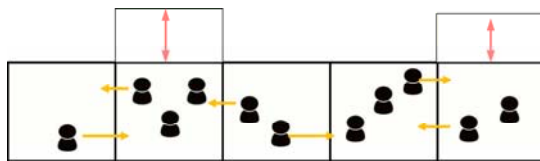


図 5 Dynamic Beach Problem Domain

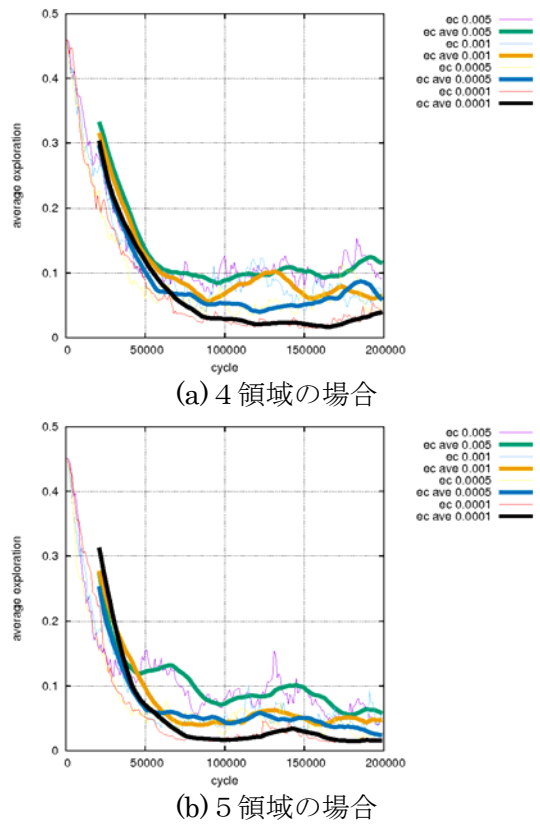


図 6 DBPD における探索率の変化

4. 2 トップダウンアプローチ

本研究では、図 1 に示すマルチエージェント資源共有問題のうち、特に大規模な人流を対象として取り上げて、トップダウンに外人流を制御する場合の手法を提案した。

数万人規模の人々が集まるイベントにおいては、運営者は来場者にイベントを楽しんでもらうことを意図するとともに、事前に誘導計画を立案して、安全で効率的なイベント運営を目指している。しかし、歩行者の誘導に関しては、道路交通における交通情報の収集・提供、信号制御といった車両を対象とした交通管制に比べて、大規模な来場者を対象とした誘導、停止および分断といった規制の効果は定量的に論じられていない。

このような背景を踏まえて、雑踏警備で行われる誘導手法のモデルを構築し、歩行者シミュレータへの実装を目指す。構築する群集規制モデルは、一次元空間モデルを採用している歩行者シミュレーション CrowdWalk への実装を行った。誘導手法を実装した歩行者シミュレータを用いることで、花火大会を始めとする屋外大規模イベントにおける誘導計画の効率性や安全性を定量的に検討可能にした。

深度センサによる門司港駅への流入量の計測結果とハンディ GPS による各帰宅動線の所要時間の計測結果からは、その待ち時間は門司港駅周辺であることと混雑状況の持続時間を把握することができる。門司港駅周辺の待ち時間は見物客が門司港駅に集中した

ために自然発生的に生じたものではなく警備会社によって行われた JR 門司港駅への入場規制によって生じている。そのため、どの帰宅動線から門司港駅への入場をさせるか、またどの程度の人を入場させるかといったオペレーションが重要であることが分かる。指定時間になれば通行可・不可を切り替える交通信号のような挙動では不十分である。

そのため、誘導指示者が観測対象となる歩行可能領域にいる見物客数を入力として、どの帰宅動線から門司港駅に何人を入場させるかを出力とする動的な誘導モデルを構築する。この動的な誘導モデルは歩行者シミュレータ CrowdWalk において、ある時刻に指定されたリンク上にいる歩行者エージェントの総和を算出し、その総和が閾値に達した場合には指定されたノードの通行可とし、その後指定されたノードの通過人数が閾値に達した場合には再び通行不可とするモデルとして実装した。この動的な誘導モデルを歩行者シミュレータ CrowdWalk に追加することで、実際には計測が難しい各見物客の門司港駅周辺での待ち時間や駅までの所要時間に与える影響を検証可能とした。

図 7 は、関門海峡花火大会の実際の帰宅動線における分断誘導モデルの実装事例を示している。実際の帰宅動線 1～3 で行われた分断誘導の位置を設定し、さらに通過禁止時刻と通過禁止停止時刻をパラメータとして与えることで、実際に行われた分断誘導を再現した。

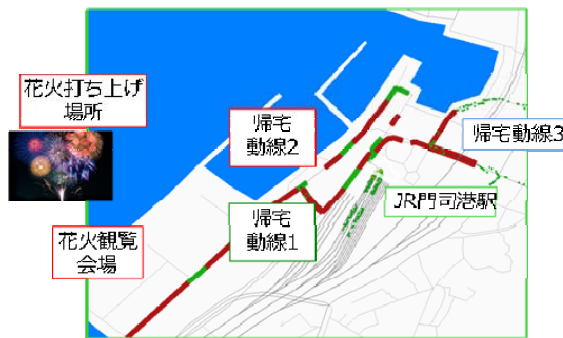


図 7 関門海峡花火大会の帰宅動線における分断誘導モデルの実装事例

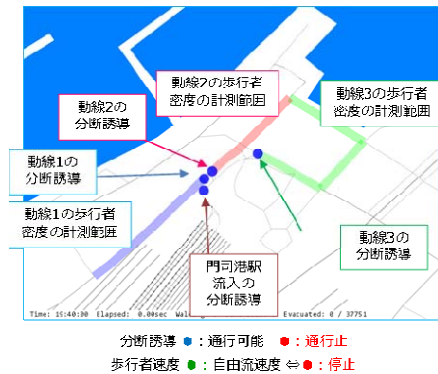


図 8 帰宅動線の制御への動的誘導モデルの事例

歩行者密度に応じた動的な分断誘導の適用事例として、門司港駅に流入する三動線の流入制御を取り上げて、下記の二つの誘導手法を比較する。

- ・誘導手法 1：静的な分断誘導
 - 門司港駅に流入する三動線断誘導箇所を順番に各 1 分ずつ通行可能にする。
- ・誘導手法 2：動的な分断誘導
 - 動線ごとの歩行者密度が計測可能という想定の中で、歩行者密度の高い動線の分断誘導を 1 分間通行可能にする。
 - ただし、全箇所が $0.2 \text{ 人}/\text{m}^2$ ならば、全箇所通行可能にする。

シミュレーション条件としては、歩行者の発生時間は、19:40-23:20 として、発生人数は合計で 37,372 人とする。帰宅動線 1 から 3 までの通過人数を帰宅動線 1 : 10,000 人、帰宅動線 2 : 7,827 人、帰宅動線 3 : 19,545 人と設定する。また、門司港駅への流入：可・不可の時間は実データを利用する。

シミュレーションの結果を 図 9、図 10 に示す。図 9 は静的な分断誘導における指定経路上の歩行者密度の推移を示している。通過人数の多い帰宅動線 3 に対して、帰宅動線 1、2 と同様の通過時間しか割り当てていないため、帰宅動線 3 の歩行者密度が上がり、 $1.2 \text{ 人}/\text{m}^2$ の状態が約 90 分間続いている。図 10 は動的な分断誘導における指定経路上の歩行者密度の推移を示している。歩行者密度の高い帰宅動線を優先して通過させるため、3 動線の歩行者密度は均衡しており $0.8 \text{ 人}/\text{m}^2$ の状態が約 50 分続いている。

本章では、人流制御のトップダウンアプローチとして歩行者密度に応じた動的な分断誘導モデルの実装を行い、門司港駅に流入する三動線の制御をおこなう事例において、静的な分断誘導と動的な分断誘導の比較を行った。

各動線の歩行者密度に対して静的な分断誘導では、来場者の特定の帰宅動線への偏りが生じた場合には対応できずに高密度の状態が続くことを確認できた。一方で、各動線の歩行者密度に対して動的な分断誘導では来場者の特定の帰宅動線の偏りに対して、各動線の密度を低く均等に保つことができることを確認した。

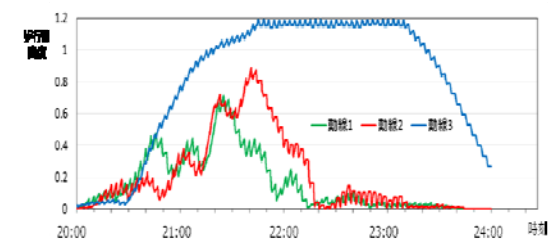


図 9 静的な分断誘導における帰宅動線中の指定経路上の歩行者密度

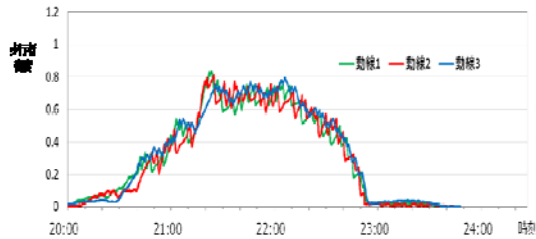


図10 動的な分断誘導における帰宅動線中の指定経路上の歩行者密度

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計2件)

- ① Takuya Okano, Itsuki Noda, "Adaptation Method of the Exploration Ratio Based on the Orientation of Equilibrium in Multi-Agent Reinforcement Learning Under Non-Stationary Environments", Journal of Advanced Computational Intelligence and Intelligent Informatics, 21, 5, pp. 939--947, Sep., 2017.
- ② 野田五十樹, 南方英明, 小林邦和, 杉浦藤虎, 武村泰範, 秋山英久, 岡田浩之, "ロボカップ西暦2050年を目指して(その1)", 知能と情報, 29, 1, pp. 2-13, 2月, 2017.

[学会発表] (計8件)

- ① 野田五十樹, "マルチエージェント社会シミュレーションの可能性", OR学会北海道支部講演会, 招待講演, OR学会, 札幌, 2月, 2018.
- ② 山下倫央, "大規模群集流動の制御に向けたシミュレーション環境の開発と応用", 複雑コミュニケーションサイエンス研究会, 8月, 2017.
- ③ 野田五十樹, "マルチエージェントシミュレーションにおけるパレートデータ同化についての考察", JAWS 2016, 岐阜羽島, 9月, 2016.
- ④ 山下倫央, 松島裕康, 野田五十樹, "屋外大規模イベントの流動制御に向けた

動的群集誘導モデルの構築", 人工知能学会全国大会 予稿集, 人工知能学会, 北九州, 6月, 2016.

- ⑤ Itsuki Noda, "RoboCup: a grand challenge toward 2050", Prof. of IROS 2016 Workshop : Challenges in Robot Competitions, IEEE-RAS, Daejeon, Korea, Oct., 2016.
- ⑥ 野田五十樹, "マルチエージェント学習とExploration率", FUN-AI 2017, 北海道・大沼, 3月, 2017.
- ⑦ 岡野拓哉, 野田五十樹, "マルチエージェント学習における探査率の進化的更新と全体最適からの乖離", SIG-DOCMAS研究会, 人工知能学会, 慶応大学, 11月, 2015.
- ⑧ 岡野拓哉, 野田五十樹, "マルチエージェント強化学習における進化的探査率調整法とその拡張", 情報処理学会第78回全国大会, 慶応大学, 3月, 2016.

[図書] (計0件)

[産業財産権]

- 出願状況 (計0件)
- 取得状況 (計0件)

[その他]

ホームページ等
(なし)

6. 研究組織

(1) 研究代表者

野田 五十樹 (NODA, Itsuki)
国立研究開発法人産業技術総合研究所・情報・人間工学領域・総括研究主幹
研究者番号: 40357744

(2) 研究分担者

山下 倫央 (YAMASHITA, Tomohisa)
北海道大学・大学院情報科学研究科・情報理工学専攻・複合情報工学講座・調和系工学研究室・准教授
研究者番号: 50415759